ON APPROXIMATIONS FOR THE DISTRIBUTIONS OBTAINED FROM MULTIPLE EVENTS

BY

H. S. STEYN [1])

(Communicated by Prof. J. F. KOKSMA at the meeting of September 29, 1962)

1. *Introduction*

In a previous paper (STEYN, 1956) the author studied a family of univariable discrete distributions of which the trinomial and univariable multinomial distributions are special cases. A simple example is the following. Consider an infinite population, each item of which belongs to one of $k+1$ classes: 0, 1, 2, ..., k.

The items of class i will cause an event E to occur i times. Clearly, the probability $f(x)$ for the event E to occur x times when n independent trials are made from such a population will be given by the coefficient of t^x in

$$G_1(t) = (p_0 + p_1 t + p_2 t^2 + \ldots + p_k t^k)^n, \tag{1}$$

where p_i is the probability for obtaining a member from the class i and $\sum_{i=0}^{k} p_i = 1$. Similarly, the univariable negative multinomial distribution, which is obtained when the trials are stopped on obtaining the mth failure (i.e. a trial in which no E occurs) is generated by the coefficient of t^x in

$$G_2(t) = p_0^m (1 - p_1 t - p_2 t^2 - \ldots - p_k t^k)^{-m}. \tag{2}$$

The moments of the univariate multinomial distribution and the negative multinomial distribution generated by (1) and (2) respectively were given in the paper mentioned above.

In the present paper certain approximations of these interesting distributions are studied for cases of one and more variables for finite and infinite populations. For the normal approximations of the distributions obtained for multiple events it is shown how the quadratic forms, which then follow χ^2 distributions, are obtained in a form similar to that derived by STEYN (1955) for the ordinary multinomial and negative multinomial distributions. Simple transformations are then derived to give new variables with first and second order moments equal to that of the corresponding ordinary multinomial and negative multinomial

[1]) This research was partly made possible by a research grant from the South African Council for Scientific and Industrial Research.

distributions. Further, the problem of approximations for multiple events in cases of finite populations is discussed and compared with the ordinary hypergeometric and negative hypergeometric distributions. Lastly, distributions which are associated with rare multiple events are compared with the well-known Poisson distribution.

2. *Transformations for χ^2 approximations*

(a) The multinomial case:

It follows from (1) that on substituting $t = e^{\alpha}$ that the cumulant generating function $\log \{G_1(e^{\alpha})\}$ will yield cumulants which are all of the order n when n is large, so that the probability function will then in standard units approximate the normal distribution. Since

$$\mu_1' \equiv E(x) = n \sum_{i=1}^{k} ip_i \quad \text{and} \quad \sigma_x^2 = n \left\{ \sum_1^k i^2 p_i - \left(\sum_1^k ip_i \right)^2 \right\} \tag{3}$$

it follows that, for n large,

$$f(x) = C. \exp \left\{ -\frac{1}{2} \frac{\left(x - n \sum_1^k ip_i \right)^2}{n \left[\sum_1^k i^2 p_i - \left(\sum_1^k ip_i \right)^2 \right]} \right\}, \quad \text{where} \tag{4}$$

C is a constant.

The equation (4) shows that

$$\eta^2 = \frac{\left(x - n \sum_1^k ip_i \right)^2}{n \left[\sum_1^k i^2 p_i - \left(\sum_1^k ip_i \right)^2 \right]}$$

is asymptotically distributed as χ^2 with one degree of freedom.

It is now easily shown by elementary algebra that:

$$\eta^2 = \frac{\left\{ \frac{\sum_1^k ip_i}{\sum_1^k i^2 p_i} x - n \frac{\left(\sum_1^k ip_i \right)^2}{\sum_1^k i^2 p_i} \right\}^2}{n \frac{\left(\sum_1^k ip_i \right)^2}{\sum_1^k i^2 p_i}} + \frac{\left\{ \left(n - \frac{\sum_1^k ip_i}{\sum_1^k i^2 p_i} x \right) - n \left(1 - \frac{\left(\sum_1^k ip_i \right)^2}{\sum_1^k i^2 p_i} \right) \right\}^2}{n \left(1 - \frac{\left(\sum_1^k ip_i \right)^2}{\sum_1^k i^2 p_i} \right)}.$$

Substituting now

$$z = \frac{\sum_1^k ip_i}{\sum_1^k i^2 p_i} x \quad \text{and} \quad z_0 = n - z,$$

we have,

$$E(z) = \frac{\sum_1^k ip_i}{\sum_1^k i^2 p_i} E(x) = n \frac{\left(\sum_1^k ip_i\right)^2}{\sum_1^k i^2 p_i},$$

so that,

$$(5) \quad \begin{cases} \eta^2 = \dfrac{\{z - E(z)\}^2}{E(z)} + \dfrac{\{z_0 - E(z_0)\}^2}{E(z_0)} \\ \quad = \dfrac{(z-np)^2}{np} + \dfrac{(z_0 - nq)^2}{nq}, \text{ where } p = \dfrac{\left(\sum_1^k ip_i\right)^2}{\sum_1^k i^2 p_i} \text{ and } q = 1 - p. \end{cases}$$

Now, remembering that for the ordinary binomial ${}^nC_x p^x q^{n-x}$, with parameter p, where x represents the number of successes in n trials, ($x_0 = n - x$ and $q = 1 - p$), that $\frac{(x-np)^2}{np} + \frac{(x_0 - nq)^2}{nq}$ is for large n approximately distributed as χ^2 with 1 degree of freedom, the similarity for large n between $z = \frac{\sum_1^k ip_i}{\sum_1^k i^2 p_i} x$, in the case of multiple events, and a binomial variable with parameters n and $p = \frac{(\sum_1^k ip_i)^2}{\sum_1^k i^2 p_i}$ is clearly seen. Clearly, z need not be an integer.

Also, $z = \frac{\sum_1^k ip_i}{\sum_1^k i^2 p_i} x$ has for any value of n, the same mean and variance as the binomial distribution with parameters n and p. This follows from

$$E(z) = n \frac{\left(\sum_1^k ip_i\right)^2}{\sum_1^k i^2 p_i} = np$$

and

$$\sigma_z^2 = \frac{\left(\sum_1^k ip_i\right)^2}{\left(\sum_1^k i^2 p_i\right)^2} \sigma_x^2 = n \frac{\left(\sum_1^k ip_i\right)^2}{\left(\sum_1^k i^2 p_i\right)^2} \left\{ \sum_1^k i^2 p_i - \left(\sum_1^k ip_i\right)^2 \right\} = npq.$$

Further, consider the case of two events; an event E which occurs either 0 or 1 or 2 ... or k times in each trial and an event F which can occur either 0 or 1 or 2 ... or h times in each trial. If further E and F are mutually exclusive, the probability generating function of the probability $f(x, y)$ for obtaining a total of x times for E and y times for F is given by the coefficient of $t^x u^y$ in:

$$(6) \qquad G(t, u) = \left(p_0 + \sum_{i=1}^k p_i t^i + \sum_{j=1}^h p_j^{(1)} u^j\right)^n,$$

where p_i is the probability for E to occur i times in each trial $i=1,2,\ldots,k$, and $p_j^{(1)}$, the probability for F to occur j times, $j=1, 2, \ldots, h$, in each trial and $p_0 = 1 - \sum_1^k p_i - \sum_1^h p_j^{(1)}$.

From (6), by putting $t=e^\alpha$ and $u=e^\beta$ and taking logarithms, it follows that all cumulants are certainly of the first degree in n, and in particular:

$$(7)\quad \begin{cases} \mu_{10}' \equiv E(x) = n \sum_{i=1}^{k} i p_i, \;\; \mu_{01}' \equiv E(y) = n \sum_{j=1}^{h} j p_j^{(1)} \\ \sigma_x^2 \equiv E(x-\mu_{10}')^2 = n \left\{ \sum_1^k i^2 p_i - \left(\sum_1^k i p_i \right)^2 \right\} \\ \sigma_y^2 \equiv E(y-\mu_{01}')^2 = n \left\{ \sum_1^h j^2 p_j^{(1)} - \left(\sum_1^h j p_j^{(1)} \right)^2 \right\} \text{ and} \\ \varrho \sigma_x \sigma_y \equiv E(x-\mu_{10}')\,(y-\mu_{01}') = -n \left(\sum_1^k i p_i \right) \left(\sum_1^h j p_j^{(1)} \right). \end{cases}$$

In standard units all cumulants higher than the 2nd order will tend to zero as n tends to infinity, so that $f(x, y)$ may be approximated by a bivariate normal distribution. For large n therefore,

$$f(x, y) \doteqdot C.\exp\{-\tfrac{1}{2}Q(x, y)\},$$

where C is a constant and

$$Q(x, y) = \frac{1}{1-\varrho^2} \left\{ \frac{(x-\mu_{10}')^2}{\sigma_x^2} - \frac{2\varrho(x-\mu_{10}')\,(y-\mu_{01}')}{\sigma_x \sigma_y} + \frac{(y-\mu_{01}')^2}{\sigma_y^2} \right\}.$$

It is well-known that $Q(x, y)$, the quadratic exponent in the bi-variate normal distribution is distributed like χ^2 with two degrees of freedom. Using the moments as given in (7) above it follows after some algebraical operations and substitution of

$$z = \frac{\sum_1^k i p_i}{\sum_1^k i^2 p_i} x, \quad w = \frac{\sum_1^h j p_j^{(1)}}{\sum_1^h j^2 p_j^{(1)}} y,$$

$$E(z) = n \frac{\left(\sum_1^k i p_i \right)^2}{\sum_1^k i^2 p_i} = np \text{ (say)} \quad \text{and} \quad E(w) = n \frac{\left(\sum_1^h j p_j^{(1)} \right)^2}{\sum_1^h j^2 p_j^{(1)}} = np^{(1)} \text{ (say)}$$

that

$$(8)\quad \begin{cases} Q(x, y) = \dfrac{(z-E(z))^2}{E(z)} + \dfrac{(w-E(w))^2}{E(w)} + \dfrac{(n-z-w-E(n-z-w))^2}{E(n-z-w)} \\ \qquad = \dfrac{(z-np)^2}{np} + \dfrac{(w-np^{(1)})^2}{np^{(1)}} + \dfrac{(n-z-w-n(1-p-p^{(1)}))^2}{n(1-p-p^{(1)})}. \end{cases}$$

The resemblance of (8) with χ^2 obtained for the case of ordinary multinomial sampling is clear (see STEYN (1955)).

Now, consider the case of r events $E_1, E_2, \ldots, E_r$ which occur only separately. If $E_{(l)}$ may occur s times, with probability $p_s^{(l)}$ $(s=1, \ldots, h_l;$ $l=1, \ldots, r)$, while p_0 is the probability that none of the $E_1, E_2, \ldots, E_r$ occur, then the transformations

$$z_i = \frac{\sum_{j=1}^{h_i} j p_j^{(i)}}{\sum_{j=1}^{h_i} j^2 p_j^{(i)}} x_i, \tag{9}$$

will have the result that,

$$\sum_{i=0}^{r} \frac{(z_i - np^{(i)})^2}{np^{(i)}}, \text{ where } p^{(i)} = \frac{\left(\sum_{j=1}^{h_i} j p_j^{(i)}\right)^2}{\sum_{j=1}^{h_i} j^2 p_j^{(i)}}, \; z_0 = n - \sum_{i=1}^{r} z_i \tag{10}$$

and

$$p^{(0)} = 1 - \sum_{i=1}^{r} p^{(i)}, \tag{11}$$

is for large n approximately distributed as χ^2 with r degrees of freedom.

The proof follows exactly as in the case of two events by considering the probability function $f(x_1, x_2, \ldots, x_r)$ generated by the coefficient of $t_1^{x_1} t_2^{x_2} \ldots t_r^{x_r}$ in

$$\left(p_0 + \sum_{j=1}^{h_1} p^{(1)} t_1^j + \sum_{j=1}^{h_2} p_j^{(2)} t_2^j + \ldots + \sum_{j=1}^{h_r} p_j^{(r)} t_r^j\right)^n. \tag{12}$$

It is easily shown that the variables z_i have the same first and second order moments as a multinomial distribution with parameters n and $p^{(i)}$, $i=1, 2, \ldots, r$.

(b) The negative multinomial case:

It follows from (2) after substituting $t=e^\alpha$ that the cumulant generating function $\log\{G_2(e^\alpha)\}$ will yield cumulants which are all of the order m when m is large. The probability function generated by (2) will, therefore, when m is large, approximate the normal distribution:

Again, using (see STEYN 1956, equation 17):

$$\mu_1' \equiv E(x) = m \sum_{i=1}^{k} i p_i/p_0 \text{ and } \sigma_x^2 = m\left\{\sum_{i=1}^{k} i^2 p_i/p_0 + \left(\sum_{i=1}^{k} i p_i/p_0\right)^2\right\} \tag{13}$$

it follows that for large m

$$f(x) \doteqdot C. \exp\left\{-\frac{1}{2} \frac{\left(x - m\sum_1^k i p_i/p_0\right)^2}{m\left[\sum_1^k i^2 p_i/p_0 + \left(\sum_1^k i p_i/p_0\right)^2\right]}\right\}, \text{ where}$$

C is a constant.

It thus follows after making the substitution

$$z = \frac{\sum_1^k ip_i}{\sum_1^k i^2 p_i} x, \text{ so that } E(z) = m \frac{\left(\sum_1^k ip_i\right)^2}{p_0 \sum_1^k i^2 p_i},$$

that

$$(14) \quad \begin{cases} \dfrac{\left(x - m\sum_1^k ip_i/p_0\right)^2}{m\left\{\sum_1^k i^2 p_i/p_0 + \left(\sum_1^k ip_i/p_0\right)^2\right\}} = \dfrac{(z-E(z))^2}{E(z)} - \dfrac{(m+z-E(m+z))^2}{E(m+z)} \\ = \dfrac{\left(z - m\dfrac{\pi}{1-\pi}\right)^2}{m\dfrac{\pi}{1-\pi}} - \dfrac{(m+z-m/(1-\pi))^2}{m/(1-\pi)} \text{ where } \dfrac{\pi}{1-\pi} = \dfrac{1}{p_0}\dfrac{\left(\sum_1^k ip_i\right)^2}{\sum_1^k i^2 p_i}, \end{cases}$$

is, for large m approximately distributed like χ^2 with one degree of freedom.

Again, the transformed variate z has for any value of m the same mean and variance as the negative binomial (or Pascal) probability function with parameters m and π. Using (13) this follows from

$$E(z) = \frac{\sum_{i=1}^k ip_i}{\sum_{i=1}^k i^2 p_i} E(x) = \frac{m}{p_0} \frac{\left(\sum_1^k ip_i\right)^2}{\sum_1^k i^2 p_i} = m\frac{\pi}{1-\pi},$$

and

$$\sigma_z^2 = \frac{\left(\sum_1^k ip_i\right)^2}{\left(\sum_1^k i^2 p_i\right)^2} \sigma_x^2 = \frac{m\left(\sum_1^k ip_i\right)^2}{\left(\sum_1^k i^2 p_i\right)^2} \left\{\sum_1^k i^2 p_i/p_0 + \left(\sum_1^k ip_i/p_0\right)^2\right\}$$

$$= m\left(\frac{\pi}{1-\pi} + \frac{\pi^2}{(1-\pi)^2}\right) \text{ or } \frac{m\pi}{(1-\pi)^2}$$

which when compared with (13) for $k=1$ gives the required result.

The case of two or more variables follows similarly. If for the r events described in section 2(a) above, the trials are stopped on obtaining the mth failure (i.e. none of the $E_1, E_2, \ldots, E_r$ occur) then using the same definitions as before, it follows from the probability generating function,

$$(15) \qquad p_0{}^m\left(1 - \sum_{j=1}^{h_1} p_j^{(1)} t_1{}^j - \sum_{j=1}^{h_2} p_j^{(2)} t_2{}^j - \ldots - \sum_{j=1}^{h_r} p_j^{(r)} t_r{}^j\right)^{-m}$$

that

$$\mu_s \equiv E(x_s) = m\sum_{j=1}^{h_s} jp_j^{(s)}/p_0, \quad \sigma_{x_s}^2 = m\left\{\sum_{j=1}^{h_s} j^2 p_j^{(s)}/p_0 + \left(\sum_{j=1}^{h_s} jp_j^{(s)}/p_0\right)^2\right\}$$

and

$$\varrho_{uv}\sigma_{x_u}\sigma_{x_v} \equiv m \sum_{i=1}^{h_u} (ip_i^{(u)}/p_0) \sum_{j=1}^{h_v} jp_j^{(v)}/p_0.$$

Next, the transformations (9) applied to the quadratic form

$$U^2 = (x-\mu)' \, V^{-1}(x-\mu),$$

where in the usual notation V is the matrix of the above variances and covariances and $(x-\mu)'$ the row vector $(x_1-\mu_1, \ldots, x_r-\mu_r)$, will yield

$$(16) \qquad \left\{ \begin{aligned} U^2 = {} & \frac{(z_1-E(z_1))^2}{E(z_1)} + \frac{(z_2-E(z_2))^2}{E(z_2)} + \ldots + \frac{(z_r-E(z_r))^2}{E(z_r)} \\ & - \frac{\left(m + \sum_{i=1}^{r} z_i - E\left(m + \sum_{i=1}^{r} z_i\right)\right)^2}{E\left(m + \sum_{i=1}^{r} z_i\right)}, \end{aligned} \right.$$

or

$$(17) \qquad U^2 = \sum_{i=1}^{r} \frac{(z_i - m\pi^{(i)}/\pi^{(0)})^2}{m\pi^{(i)}/\pi^{(0)}} - \frac{\left(m + \sum_{1}^{r} z_i - m/\pi^{(0)}\right)^2}{m/\pi^{(0)}},$$

where

$$\frac{\pi^{(i)}}{\pi^{(0)}} = \frac{1}{p_0} \frac{\left(\sum_{1}^{h_i} jp_j^{(i)}\right)^2}{\sum_{1}^{h_i} j^2 p_j^{(i)}} \quad \text{for } i = 1, \ldots, r \text{ and } \pi^{(0)} = 1 - \sum_{1}^{r} \pi^{(i)}.$$

Now, U^2 is for large m, approximately distributed as χ^2 with r degrees of freedom. The expression (17) may be compared with the result obtained by STEYN (1955, page 595), for ordinary negative multinomial sampling.

(c) The case of finite populations:

(i) *The factorial multinomial case.*

Consider a finite population of size N of which Np_i items will cause an event E to occur i times, $i = 0, 1, \ldots, k$. For sampling without replacement from such a population it was shown by STEYN (1956, p. 194) that the probability $f(x)$ for having a total of x successes in n trials is given by the sum of those terms for which $\sum_1^k ir_i = x$ in the factorial power expansion of

$$\frac{1}{N^{!n}} \left(\sum_{i=1}^{k} Np_i\right)^{!n} = \frac{1}{N^{!n}} \sum_{r_1, \ldots, r_k} \frac{n!\,(Np_0)^{!(n-r)}}{(n-r)!} \prod_{i=1}^{k} \frac{(Np_i)^{!r_i}}{r_i!},$$

where $\sum_{i=1}^{k} r_i = r$. The mean of this distribution is the same as that given in (3) while the variance differs from the variance given in (3) only by

a constant factor $(N-n)(N-1)^{-1}$. It, therefore, follows as in (5), that for large values of the parameters N and n,

$$\eta^2 = \frac{N-1}{N-n}\left\{\frac{(z-np)^2}{np} + \frac{(z_0-nq)^2}{nq}\right\}, \tag{18}$$

where z, z_0 and p are defined as in (5), is approximately distributed as χ^2 with one degree of freedom.

Comparing this with the expression for χ^2 in the ordinary case of a hypergeometric or factorial binomial distribution (see STEYN 1955, p. 595) the similarity between the variable $z = \frac{\sum_1^k ip_i}{\sum_1^k i^2p_i}x$, for large values of N and n, and a hypergeometric variable is clearly seen. Also, the mean and variance of z is the same as that of a hypergeometric distribution (or factorial binomial) with parameters N, n and

$$p = \frac{\left(\sum_1^k ip_i\right)^2}{\sum_1^k i^2p_i}, \text{ since,}$$

$$E(z) = \frac{\sum_1^k ip_i}{\sum_1^k i^2p_i}E(x) = n\frac{\left(\sum_1^k ip_i\right)^2}{\sum_1^k i^2p_i} = np,$$

and

$$\sigma_z^2 = \frac{\left(\sum_1^k ip_i\right)^2}{\left(\sum_1^k i^2p_i\right)^2}\sigma_x^2 = \frac{N-n}{N-1}n\frac{\left(\sum_1^k ip_i\right)^2}{\left(\sum_1^k i^2p_i\right)^2}\left\{\sum_1^k i^2p_i - \left(\sum_1^k ip_i\right)^2\right\} = \frac{N-n}{N-1}npq.$$

Similarly, for the finite populations corresponding to the multinomial cases for more than one event (or variables), where the $p_j^{(i)}$ is the fraction of the items of class j in the ith subpopulation (i.e. causing event E_i to happen j times), it may easily be verified that the means for the finite populations are the same as those already obtained for the corresponding infinite populations while the variances and product moments (corresponding to (7)) will differ only by a factor $(N-n)(N-1)^{-1}$, which shows that the expression (10) must then be multiplied by the factor $(N-1)/(N-n)$.

(ii) *The Negative Factorial Multinomial case.*

When sampling without replacement from the finite population mentioned above is stopped on obtaining the mth failure the cases corresponding to 2(b) above will differ in only two aspects. This can best be seen by writing down the mean and variance for the negative factorial

multinomial distribution corresponding to the expression (2) and which is given in the previous mentioned paper (STEYN 1956, p. 195 – using m instead of $m+1$) as:

$$\mu_1' = m \sum_{i=1}^{k} iNp_i/(Np_0+1)$$

and

$$\sigma_x^2 = \frac{(Np_0-m-1)}{(Np_0+2)}\, m \left\{ \left(\sum_{i=1}^{k} iNp_i/(Np_0+1) \right)^2 + \sum_{i=1}^{k} i^2 Np_i/(Np_0+1) \right\}.$$

Comparing these moments with those in (13) it is seen that p_i/p_0 is now replaced by $Np_i/(Np_0+1)$ and a factor $(Np_0-m-1)/(Np_0+2)$ further appears, so that from (14) it follows that:

$$\frac{(Np_0+2)}{(Np_0-m-1)} \left\{ \frac{(z-E(z))^2}{E(z)} - \frac{(m+z-E(m+z))^2}{E(m+z)} \right\},$$

where

$$z = \frac{\sum_{i=1}^{k} ip_i}{\sum_{i=1}^{k} i^2 p_i}\, x \quad \text{and} \quad E(z) = \frac{mN}{Np_0+1} \frac{\left(\sum_{i=1}^{k} ip_i \right)^2}{\sum_{i=1}^{k} i^2 p_i},$$

is, for large values of N and m, distributed as χ^2 with one degree of freedom. Also, this transformed variate z has the same mean and variance as the negative factorial binomial (or inverse hypergeometric) distribution with parameters N, m and π where

$$\frac{N\pi}{N(1-\pi)+1} = \frac{\left(\sum_{i=1}^{k} iNp_i \right)^2}{(Np_0+1)\left(\sum_{1}^{k} i^2 Np_i \right)}.$$

Also, for the case of finite populations the expression corresponding to (16) will now be:

$$(19) \qquad \frac{Np_0+2}{Np_0-m-1} \left\{ \sum_{1}^{r} \frac{(z_i-E(z_i))^2}{E(z_i)} - \frac{\left(m + \sum_{i=1}^{r} z_i - E\left(m + \sum_{i=1}^{r} z_i \right) \right)^2}{E\left(m + \sum_{i=1}^{r} z_i \right)} \right\},$$

where the z_i is defined in (9) and

$$E(z_i) = \frac{mN}{Np_0+1} \frac{\left(\sum_{j=1}^{h_i} jp_j^{(i)} \right)^2}{\sum_{j=1}^{h_i} j^2 p_j^{(i)}}.$$

This quadratic form is for large N and m approximately distributed like χ^2 with r degrees of freedom.

3. *Limiting Forms for Rare Multiple Events*

When substituting $p_0 = 1 - \sum_1^k p_i$ in the generating function (1), it follows that

$$G_1(t) = \{1 + p_1(t-1) + p_2(t^2-1) + \ldots + p_k(t^k-1)\}^n,$$

such that, when n is large but p_i of $0(1/n)$, $i = 1, 2, \ldots, k$, $G_1(t)$ may be approximated by

$$(20) \qquad e^{\theta_1(t-1)+\theta_2(t^2-1)+\ldots+\theta_k(t^k-1)} \text{ or } e^{-\sum_1^k \theta_i} e^{\sum_1^k \theta_i t^i}, \text{ where } \theta_i = np_i.$$

For $k=2$, the probability function given by the coefficient of t^x in (20) will clearly be:

$$e^{-(\theta_1+\theta_2)} \left\{ \frac{\theta_1{}^x}{x!} + \frac{\theta_1{}^{x-2}}{(x-2)!}\frac{\theta_2}{1!} + \frac{\theta_1{}^{x-4}}{(x-4)!}\frac{\theta_2{}^2}{2!} + \ldots \right\}.$$

By putting $t = e^\alpha$ in (20) and taking the logarithms, the first three cumulants follow immediately as:

$$(21) \qquad k_1 \equiv \mu_1' = \sum_1^k i\theta_i, \quad k_2 \equiv \sigma^2 = \sum_1^k i^2\theta_i,$$

and $k_3 \equiv \mu_3$ (the central third moment) $= \sum_1^k i^3\theta_i$. Further, all higher cumulants are of the same order as the θ_i's.

Thus, for large values of the θ_i's, when using standard units,

$$\left(x - \sum_1^k i\theta_i\right) \Big/ \left(\sum_1^k i^2\theta_i\right)^{\frac{1}{2}}$$

is approximately normally distributed, so that, using

$$(22) \qquad z = \frac{\sum_1^k i\theta_i}{\sum_1^k i^2\theta_i} x, \text{ it follows that } \frac{(z-\lambda)^2}{\lambda}, \text{ where } \lambda = \frac{\left(\sum_1^k i\theta_i\right)^2}{\sum_1^k i^2\theta_i}$$

is distributed as χ^2 with one degree of freedom.

It is easily seen from (21) that,

$$E(z) = \frac{\sum_1^k i\theta_i}{\sum_1^k i^2\theta_i} E(x) = \lambda, \quad (\lambda > 0)$$

and

$$\sigma_z{}^2 = \frac{\left(\sum_1^k i\theta_i\right)^2}{\left(\sum_1^k i^2\theta_i\right)^2} \sum_1^k i^2\theta_i = \lambda.$$

This χ^2-approximation as well as the first two moments of the distribution of $z = \frac{\sum_1^k i\theta_i}{\sum_1^k i^2\theta_i} x$ show that it is worthwhile to investigate the correspondence between the distribution of z and a Poisson distribution with parameter $\lambda = \frac{(\sum_{i=1}^k i\theta_i)^2}{\sum_{i=1}^k i^2\theta_i}$. Clearly, from (21), the central third moment of z is

$$\mu_3 = \frac{\left(\sum_{i=1}^k i\theta_i\right)^3}{\left(\sum_{i=1}^k i^2\theta_i\right)^3} \cdot \sum_{i=1}^k i^3\theta_i$$

$$= \lambda \left\{ \left(\sum_{i=1}^k i\theta_i \cdot \sum_{i=1}^k i^3\theta_i \right) \Big/ \left(\sum_{i=1}^k i^2\theta_i \right)^2 \right\}$$

$$= \lambda\{1+\Theta\}, \text{ where } \Theta = \frac{\sum_{i=1}^k i\theta_i \sum_{i=1}^k i^3\theta_i - \left(\sum_{i=1}^k i^2\theta_i\right)^2}{\left(\sum_{i=1}^k i^2\theta_i\right)^2}.$$

The numerator of Θ is equal to

$$\sum_{i<j} (ij^3 + i^3j - 2i^2j^2)\,\theta_i\theta_j = \sum_{i<j} ij(j-i)^2\,\theta_i\theta_j.$$

The central third moment of z, therefore, differs from that of a Poisson distribution with parameter λ by the factor $(1+\Theta)$.

It is clear that $\Theta > 0$ and that the numerator will contain each θ_i only in the first power, so that when any *one* of the θ_i's is large, then Θ will be small. Next, if for $i = 1, 2, \ldots, k$ we have $N \leqslant \theta_i \leqslant M$, then

$$\Theta \leqslant \frac{M^2}{N^2} \left\{ \left(\sum_1^k i \right) \left(\sum_1^k i^3 \right) - \left(\sum_1^k i^2 \right)^2 \right\} \Big/ \left(\sum_1^k i^2 \right)^2, \text{ or}$$

$$\Theta \leqslant \frac{M^2}{N^2} \{[\tfrac{1}{2}k\,(k+1)]^3 - [\tfrac{1}{6}k\,(k+1)(2k+1)]^2\} / [\tfrac{1}{6}k\,(k+1)(2k+1)]^2,$$

so that for,

(i) $k=1$, $\Theta = 0$ (as expected),

(ii) $k=2$, $\Theta \leqslant M^2/N^2 \cdot 2/25$,

(iii) $k=3$, $\Theta \leqslant M^2/N^2 \cdot 10/98$,

(iv) when $k \to \infty$, $\Theta \leqslant M^2/N^2 \cdot 1/8$, the upper limit for any k.

This shows, that if the θ_i's do not vary considerably, the Poisson approximation will give a fair approximation also for the third moment.

There is another way of approximating the probability function generated by (20) in terms of the Poisson distribution when the θ_i's are large. From (20) follows for $t = 1+\alpha$ the factorial moment generating function

$$e^{\sum_1^k i\theta_i} \cdot \prod_{i=s}^k e^{\sum_{i=s}^k {}^iC_s \theta_i \alpha^s},$$

so that $\mu'_{(r)}$, the rth factorial moment of the probability function generated by (20) is given by

$$\mu'_{(r)} = \left(\sum_1^k i\theta_i\right)^r + \frac{r(r-1)}{2!}\left(\sum_1^k i\theta_i\right)^{r-2}\left(\sum_2^k i(i-1)\theta_i\right)$$
$$+ \text{ terms of degree } r-3 \text{ and lower in the } \theta_i\text{'s.}$$
$$= \left(\sum_1^k i\theta_i\right)^r \left\{1 + \frac{\frac{r(r-1)}{2}\sum_2^k i(i-1)\theta_i}{\left(\sum_1^k i\theta_i\right)^2} + \ldots\right\} \to \left(\sum_1^k i\theta_i\right)^r \text{ when } \sum_1^k i\theta_i \to \infty.$$

Thus $\mu'_{(r)}$ tends to the factorial moment of a Poisson distribution with parameter $\sum_1^k i\theta_i$. Again, $z = \frac{\sum_1^k i\theta_i}{\sum_1^k i^2\theta_i} x$, will then for large values of the θ_i's approximately have the moments of a Poisson distribution with parameter $\lambda = \frac{(\sum_1^k i\theta_i)^2}{\sum_1^k i^2\theta_i}$.

The further extensions to more variables of the limiting forms for rare multiple events are straightforward. For example, the expression (10) will become:

$$\sum_{i=1}^r \frac{(z_i - \lambda^{(i)})^2}{\lambda^{(i)}}, \text{ where } \theta^{(i)} = np^{(i)},$$

$$z_i = \frac{\left(\sum_{j=1}^{h_i} j\theta_j{}^{(i)}\right)}{\sum_{j=1}^{h_i} j^2\theta_j{}^{(i)}} x_i \text{ and } \lambda^{(i)} = \frac{\left(\sum_{j=1}^{h_i} j\theta_j{}^{(i)}\right)^2}{\sum_{j=1}^{h_i} j^2\theta_j{}^{(i)}}.$$

Also, the limiting cases for negative multinomials need no further investigations as they are clearly the same as those for the positive multinomials. To see this we note that the expression for the generating function (2) when m is large, but p_i of $0(1/m)$, $i = 1, 2, \ldots, k$, approximates:

$$\lim_{m\to\infty}\left(1 - \sum_1^k p_i\right)^m \left(1 - \sum_1^k p_i t^i\right)^{-m} = e^{-\sum_1^k \theta_i} \cdot e^{\sum_1^k \theta_i t^i}, \text{ where } \theta_i = mp_i,$$

which is the same as (20).

The author wishes to thank Mrs. G. PRETORIUS for technical assistance while drawing up this paper.

University of South Africa, Pretoria

LITERATURE

STEYN, H. S., On Discrete Multivariate Probability Functions of Hypergeometric Type. (Koninkl. Nederl. Akademie van Wetenschappen, Proc., Series A, no. 5 (1955)).

———, On the Univariable Series $F(t) = F(a; b_1, b_2, \ldots, b_k; c; t, t^2, \ldots, t^k)$ and its Applications in Probability Theory. (Koninkl. Nederl. Akademie van Wetenschappen, Proc., Series A, **59**, no. 2 (1956)).